DOCUMENT RESUME

ED 360 988 IR 054 676

AUTHOR Van Houweling, Douglas E.; McGill, Michael J.
TITLE The Evolving National Information Network: Background and Challenges. A Report of the Technology Assessment Advisory Committee to the Commission on Preservation and Access.
INSTITUTION Commission on Preservation and Access, Washington, DC.
PUB DATE Jul 93
NOTE 37p.
AVAILABLE FROM Commission on Preservation and Access, 1400 16th Street, N.W., Suite 740, Washington, DC 20036-2217.
PUB TYPE Information Analyses (070) -- Reports - Evaluative/Feasibility (142)

EDRS PRICE MF01 Plus Postage. PC Not Available from EDRS.
DESCRIPTORS *Access to Information; *Computer Networks; Futures (of Society); Higher Education; *Information Networks; Information Utilization; Knowledge Level; Policy Formation; Research Methodology; *Technological Advancement; *Telecommunications
IDENTIFIERS Digital Transmission Systems; Historical Background; *Information Infrastructure; *Information Policy

ABSTRACT

The rapidly developing and changing networking and telecommunications environment now being implemented in the United States and across the globe is explored. The creation of a flexible and inexpensive digital network allowing instantaneous access by any individual to information of any type is now within our grasp. A primer on the technology underlying the current network environment is offered, and the status of higher education and research networking environments is described. Some of the lessons learned to date are highlighted, and the impacts of this technology on the issues of knowledge preservation and access are considered. Scenarios of possible futures using the interlocking technologies of communications, processing power, and storage are provided. Prospects and challenges in the further development of the network are also reviewed. Policy issues that cannot be ignored are those of: (1) access; (2) appropriate use; (3) funding; (4) governance; (5) security; and (6) privacy and free speech. The success of current networks is leading to the National Information Infrastructure, which will have immense impact on society. An appendix contains a brief history of national networks. (SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 360 988

The Commission on Preservation and Access

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☐ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy

The Evolving National Information Network: Background and Challenges

A Report of the Technology Assessment Advisory Committee
to the
Commission on Preservation and Access

by
Douglas E. Van Houweling
Vice Provost for Information Technology

Michael J. McGill
Director, Network Systems

University of Michigan
Ann Arbor, Michigan 48109

July 1993

"PERMISSION TO REPRODUCE THIS MATERIAL IN OTHER THAN PAPER COPY HAS BEEN GRANTED BY

M.K. Sitts

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)"

1205476

1400 16th Street, N.W., Suite 740, Washington, D.C. 20036-2217 • (202) 939-3400

A private, nonprofit organization acting on behalf of the nation's libraries, archives, and universities to develop and encourage collaborative strategies for preserving and providing access to the accumulated human record.

BEST COPY AVAILABLE

The Evolving National Information Network: Background and Challenges

A Report of the Technology Assessment Advisory Committee
to the
Commission on Preservation and Access

by
Douglas E. Van Houweling
Vice Provost for Information Technology

Michael J. McGill
Director, Network Systems

University of Michigan
Ann Arbor, Michigan 48109

July 1993

COMMITTEE PREFACE

This report completes a cycle of reports on technological developments, trends and issues pertaining to the preservation copies of print, film, analog non-print and digital materials initiated by the Technology Assessment Advisory Committee (TAAC) to the Commission on Preservation and Access. The focus of the prior reports was the changing technologies, options and interrelationships of preservation, namely micrographics and digital imaging and storage. The most recent report went beyond the preservation of our disintegrating print collection to explore the growing and likewise alarming implications of the relatively short life cycle of electronic publications, records and files as well as non-print analog materials.

This study focuses on one aspect of another essential objective of the Commission, namely, ready and universal access to the materials that are being preserved for posterity and for which the originals are rapidly becoming inaccessible.

The developing and expanding regional, national and global digital highways reflect the changing combination of communications and processing technologies. They are providing remote access to digitally stored materials and promise to provide relatively universal, convenient, cost effective and prompt access on a scale previously unrealizable. This will have a profound impact on the electronic surrogates of print and film materials, including preservation copy. Indeed, if appropriate mechanisms and collaborative efforts are established, preservation surrogates of materials will become more readily and universally accessible than either print originals or microform copies.

The author has faced a very daunting task in providing a comprehensive, yet succinct and readable, overview of a combination of technologies. We are indebted to our fellow TAAC member and principal author, Douglas van Houweling, for undertaking this ambitious task and executing it with such clarity and understanding. I also want to express appreciation to the other members of TAAC and Doug's colleagues at the University of Michigan for their comments and contributions.

--Rowland C.W. Brown
Chair, Technology Assessment Advisory Committee

Committee members are: (Chair) Rowland C.W. Brown, Consultant; Douglas van Houweling, Vice Provost for Information Technology, University of Michigan; Michael Lesk, Division Manager, Computer Science Research, Bellcore; Peter Lyman, Librarian and Dean, University of Southern California; M. Stuart Lynn, Vice President, Information Technologies, Cornell University; Robert Spinrad, Vice President, Technology Analysis and Development, Xerox Corporation; and Robert L. Street, Vice Provost and Dean of Libraries and Information Resources, Stanford University.

Published by
The Commission on Preservation and Access
1400 16th Street, NW, Suite 740
Washington, DC 20036-2217

July 1993

Reports issued by the Commission on Preservation and Access are intended to stimulate thought and discussion. They do not necessarily reflect the views of Commission members.

Additional copies are available from the above address for $15.00. Orders must be prepaid, with checks made payable to "The Commission on Preservation and Access," with payment in U.S. funds.

This paper has been submitted to the ERIC Clearinghouse on Information Resources.

The paper in this publication meets the minimum requirements of the American National Standard for Information Sciences-Permanence of Paper for Printed Library Materials ANSI Z39.48-1992.

COPYRIGHT 1993 by the Commission on Preservation and Access. No part of this publication may be reproduced or transcribed in any form without permission of the publisher. Requests for reproduction for noncommercial purposes, including educational advancement, private study, or research will be granted. Full credit must be given to the author(s) and The Commission on Preservation and Access.

Table of Contents

Introduction 1
- Overall structure of the paper 1
- Pace of change 1

Technology Primer 3
- Digitization 3
- Analog circuits and modems vs. digital circuits 3
- Line speeds 4
- Circuit switching vs. packet switching 4
- Types of packet networks 5
- Protocol standards 6
- Applications 12
- Figure 1 13
- Support services 15
- Integrity, security, and privacy 16

Current Status 17
- Technology 17
- Organization 17

What We Have Learned 19
- Requirements 19
- Rapid growth underway 21
- The need for a broadly cooperative effort 21
- The future 21

Policy Issues 21
- Access 21
- Appropriate use 22
- Funding 22
- Governance 22
- Security 22
- Privacy and free speech 23

Potential Impact 23
- The document life cycle 23
- Preservation 26
- Universities as a human knowledge repository 26
- Broad knowledge access 27
- The university 28

Conclusion 28

Appendix -- Brief History 31
- Introduction 31
- ARPANet 31
- CSNET 31
- BITNET 31
- NSFNET 32
- Library networks 32

Introduction

This document addresses the rapidly developing and changing networking and telecommunications environment now being implemented in the United States and across the globe. The creation of a flexible, relatively inexpensive digital network operating at speeds capable of supporting instantaneous access by any individual to information of any type wherever located is a prospect now within our grasp. This new environment for information transmission has the potential to revolutionize the way we publish, access, and preserve information. It has the capability to transform the way libraries acquire and preserve information and introduces new prospects for information preservation and access.

Overall structure of the paper

Evaluating the implications of the networked future for preservation of and access to information requires a broad grasp of the enabling technology. To provide this foundation of understanding, the paper first offers a primer on the technology underlying the current network environment. It then describes the status of the higher education and research networking environment, provides a summary of some of the lessons that we have learned from experience with these networks , and suggests impacts of this technology on the issues of knowledge preservation and access. It provides scenarios of possible futures employing the interlocking technologies of communications, processing power and storage. It addresses the prospects and challenges we face in the further development of the network and its use. It concludes with an outline of the policy issues we face.[1]

Pace of change

Perhaps the most important introductory comment to this paper has to do with the extraordinary pace of change in the technology underlying today's information world.[2] This assessment of telecommunications and network technology is undertaken in an atmosphere of explosive change and rapidly developing potential. The resulting opportunities for scholarship and its supporting information infrastructure are extraordinary and many of them are certainly not imaginable at present.

---

[1]This paper focuses on the network, but the network is itself imbedded in the larger organizational structure of the standards-based computing world. The international information technology community is emerging as a non-hierarchical and rapidly evolving set of inter-dependent organizations based on a commonly held set of values about open architectures, distributed management, and the importance of standards.

[2]Written during the spring of 1993, this report can, at best, provide only a snapshot of an extraordinarily rapidly changing environment. While background is provided and some prognostications are attempted, the background will soon become ancient history and the prognostications will rapidly be supplanted by the future reality.

Communications

The confluence of microelectronics and new communications media such as fiber optics, laser amplifiers, ultra high frequency and microwave radio technology; the increasingly entrepreneurial and less regulated telecommunications industry; and the global impacts of competition and increasing standards development have all led to a dramatic improvement in the cost effectiveness of telecommunications technology.

Processor power

Accompanying this revolution in technology for communications has been a more rapid growth of processing power, especially personal computers and computers embedded in other technology. This rapid increase in processing power not only allows the use of much more complicated communications protocols than ever before, but also enables strikingly increased flexibility and responsiveness in the interface between the human user and the computer. Furthermore, dramatic reductions in power consumption and size are leading to increasing possibilities of portability and ever larger capacities for manipulation of information environments. Networks have become more crucial since microprocessors are the most cost-effective source of this power and many applications depend on resources beyond the user's own machine.

Storage

Finally, striking advances in storage technology have accompanied these advances in communications technology and processing. While optical storage technology has received much attention, the advances that have had the most impact continue to be those in magnetic storage technology. It is now possible to store the contents of more than one hundred books in about nine cubic inches.

Furthermore, all the evidence is that there will be a continuing technology explosion in all of these fields -- communications, processing power and storage -- for at least two decades to come. If one can imagine it and if it will be useful, the fact that an application is not currently feasible is only a temporal statement and quite likely to be overcome by technology in the near future. As an example, we could not imagine thirty years ago that computers would become so powerful, affordable, and compact that they would effectively displace pens, pencils, and typewriters as the primary technology by which many of us would write. Today we can barely imagine that a portable networked computer tablet will displace our note pad, newspaper, magazine, book, television, dictating machine, camera, and telephone. Such an environment is by no means unlikely by the year 2010, less than twenty years from now.

Technology Primer

Digitization

Digitization of information is the single fundamental theme underlying modern data communications networks and their usefulness. Images, speech, music, diagrams, and the written word can all be translated into a sequence of numbers. That sequence can be stored, processed, and/or transmitted. At the other end of the process, the numbers can be interpreted to recreate information in the form in which it was originally expressed. Since computers are able to manipulate digitized information, advances in computer technology are now applicable to information of all types and all types of information can be transported on data communication networks.

Analog circuits and modems vs. digital circuits

Most of the communications infrastructure in the United States and around the world was designed and built primarily for the transmission of sound. Today, the networks are expected to transmit a great deal more than sound. When there was only sound to transmit it was easy to use an analog transmission circuit. This type of transmission sends information as a continuous signal with the frequency of the signal changing as the tone of the communication changed. This type of signal models speech and was well suited for voice transmission. Analog circuits can be used to transmit digital information but require that the digital information be changed to an analog signal, communicated through the voice telecommunications network and then converted back to digital information. The device that changes the signals from digital to analog and then back to digital is called a MODulator/DEModulator or MODEM.

The steadily improving capabilities of digital technology led to increasing use of digital circuits in computer communications and subsequently in communication networks. At first, digital circuits were most cost effective for connecting the major hubs of communication networks. Those hubs needed to send large amounts of information at very high speed across significant distances. Digital technologies made more efficient use of the expensive long distance facilities.

As advances in microelectronics continued, digital technologies became increasingly cost effective, and a larger and larger proportion of the communications network has been digitized. Today all of the long distance communications carriers are moving toward completely digital long distance networks. Microelectronics are now being used to dramatically reduce the cost of circuits, making digital circuits cost effective for increasingly short distances.

One force leading to the increased digitization of the network is an increasing demand to directly connect computers at different locations. While digital communication is now routine in the long distance environment and easily obtained by large organizations, it is not yet provided routinely in the small business and residential environment. Most users of personal computers in homes and small

businesses still use modems to communicate across analog circuits with other computers. Even so, the computers with which they communicate are usually connected to other computers via a digital network. A new data communication offering known as Integrated Services Digital Network (ISDN), is just now being made available in some parts of the country. The lowest speed implementation of this service, Basic Rate ISDN, has the capability of providing direct digital connections at affordable prices for home and small business use. Since the service has not been broadly and inexpensively tariffed for individual lines, it has not yet gained sufficient market volume to be inexpensively integrated into the hardware and software of personal computers. As a result, modems continue to be the dominant means of connecting homes and small businesses to the network.

Line speeds

At the same time that the analog transmission environment is being transformed to a digital transmission environment, the available transmission speeds have increased dramatically. In the 1970's, the speed of digital transmission over the analog circuits then routinely available was fifteen characters per second. Toward the end of the 1970's, it became straightforward to acquire modems that had a capability of transmitting up to 120 characters per second on a local telephone line.

Within the last few years, it has become routine for those transmissions to take place at rates of 240, 480, 960, and 1,440 characters per second. Modem prices are declining every month so that it is now possible to use a local analog telephone line to transmit information at multiple thousands of characters per second. Even so, these speeds, which can be obtained through the use of modems on analog transmission facilities, are far short of what can be obtained through digital facilities. It has long been possible to obtain, at a reasonable price, digital facilities which transmit information at 6,000 characters per second and only slightly more expensive facilities which are capable of transmitting 200,000 and 1 million characters per second. In the last year, systems capable of transmitting 5.6 million characters per second through digital facilities were beginning to appear. In other words, the analog networks and modems of the early 70's would have required more than an hour to transmit the report you're reading, while today's digital network can do it in a fraction of a second!

Circuit switching vs. packet switching

As these new digital facilities have developed, there has been a parallel set of developments in the switching required to route information from source to destination. Prior to the availability of low cost computer technology, connection between an analog telephone at one location and another at a distance was accomplished through a complicated switching system which had the capability of connecting a local pair of wires to a remote pair of wires. To be specific, twenty years ago when you dialed your telephone the dial directly actuated a set of switches which made physical connections between the wires attached to your telephone and the wires attached to the telephone you were dialing. Since this created an electrical circuit between one location and another, this switching technology became known as circuit switching.

Today, when you enter code numbers on a touch tone pad on your telephone, you are signaling a computer-based switch. This switch transmits electrical signals from your telephone by digital technology through additional switches that recreate those electrical signals at a remote telephone. In most cases, there is no direct circuit connection between your telephone and the remote telephone. In fact, once the sound coming from your telephone is transformed into digits, those digits are interspersed with other streams of digits coming from other telephones according to a carefully worked out multiplexing scheme that enables a single circuit using digital technology to carry a number of conversations simultaneously.

In contrast to voice communications, which require a relatively constant rate of information flow from one location to another, computer communications tend to require more episodic or "peaky" information flow. That characteristic, combined with the increasing availability of higher speed data communication circuits, led in the 1960's to the development of a new switching technology called packet switching. Unlike circuit switching where a dedicated connection is established between the source and the destination of information, packet switching relies on capturing blocks (packets) of information from the source, and attaching a destination address to each packet of information. The information is then forwarded to a computer which reads the destination address and determines the appropriate route for that information to follow toward its destination. The computer then sends the packet on to the next computer along the route, which repeats the process. This forwarding continues until the information is delivered to its destination. In most cases, a packet of information will traverse several such computers, called packet switches, before arriving at its destination. Depending on the amount of traffic that needs to be transmitted between any two locations, successive packets may be sent through the interconnected packet switches using different routes in response to circuit congestion or availability. Packet switching technology was initially designed to provide a robust communication structure that could survive a situation in which a large number of circuits became unavailable due to some natural or man made disaster. For the transmission of data, packet switching now dominates circuit switching and is the foundation for most of today's data communication infrastructure.

Types of packet networks

Local Area Networks and Departmental Networks

There are today two different types of data networks for local use called Local Area Networks (LANs). The first type uses a distributed architecture with a packet switch in each connected device. The second type concentrates the traffic in a number of switches considerably smaller than the number of attached computers.

Local area networks such as Ethernet and IBM's token ring all use distributed packet routing technologies through a common communications infrastructure. In these networks, all traffic from all connected machines flows over a shared wiring structure. The shared structure must have adequate capacity to pass the total traffic in the network and, in principle, allows any machine to receive any traffic, whether or not the traffic was intended for that machine.

Local area networks such as Starlan have central packet switching points that operate in cooperation with special devices in the computers attached to these networks. Only the traffic to or from a single computer flows over the wiring for any particular machine. As a result, the speed of the network only needs to be adequate to serve that machine, and traffic addressed to that machine is not accessible by other attached machines.

Campus or Corporate Networks

Networks that cover larger geographic areas such as a large campus or corporation often include some hierarchy in which local area networks are interconnected to a higher capacity network that provides connectivity to the overall campus or enterprise. In this situation, the campus network really becomes a network of networks, or an 'internet'. The large capacity network that interconnects the local area networks is often referred to as the backbone. The backbone is made up of high capacity cabling, often fiber, and packet switches called routers that connect the local area networks to the backbone.

Metropolitan area networks (MANs)

The campus network approach can be expanded to include whole metropolitan areas. These networks are becoming major users of digital packet technology serving the needs of large numbers of customers.

Wide area networks (WANs)

The most geographically extensive networks are called wide area networks (WANs), and typically use a number of packet switches to route information across the world. There are a number of commercial value added common carriers that provide these services including Sprintnet, Tymnet and Autonet. In addition, there are special purpose packet networks that have been developed to support particular applications such as the networks that support CompuServe, MCIMail and Prodigy.

Also in the category of wide area networks are the intercampus and interestablishment portions of networks such as NSFNET, BITNET, and CSNET. These latter networks have substantial impacts on the higher education and research community and are the subject of much of the discussion later in this paper. These networks now have several levels of hierarchy and interconnect many different networks. They together have become known as the Internet.

Protocol standards

In much the same way that early telephone systems evolved from a hodgepodge of independent companies, each interconnecting only telephones owned by them, to a group of cooperating interconnected companies, the data networking environment has evolved. In the early days of telephony, telephone users were often forced to have four or five different telephones on their desk so they would be able to communicate with individuals on each of the telephone systems that served other individuals with whom they wished to communicate. As the environment evolved, they were able to acquire service from one telephone company which, in turn, connected with other telephone companies, providing the appearance for the user of

a single telephone company and network which provided universal connectivity. Underlying that development were a set of agreements among the various telephone companies regarding standards for interconnection of their systems.

Similarly, standards have been developed to allow interconnection of networks[3]. A number of organizations have been instrumental in the work required to develop and codify the standards. In particular, the Institute for Electrical and Electronic Engineers, the American National Standards Institute and the National Information Standards Organization play the dominant standards roles in the United States. The International Organization for Standardization (ISO) and the International Consultative Committee for Telephony and Telegraphy (CCITT) play the dominant roles on the international scene. The standards promulgated by these agencies are usually assigned codes, which will be cited along with the names of the standards in the following discussion. Protocols are defined in detail by complementary standards, usually called protocol suites. There are two dominant protocol suites.

Transmission Control Protocol/Internet Protocol (TCP/IP)

The TCP/IP protocol suite is also known as the Department of Defense (DOD), Defense Advanced Research Projects Agency (DARPA), or Internet protocol suite. It was proposed and developed under the leadership of DARPA and has subsequently become a Department of Defense standard. TCP/IP protocols are the foundation of several thousand connected networks that together form the TCP/IP Internet. These networks form the backbone for data communications in higher education and research throughout the United States today. The TCP/IP protocol suite is very widely adopted in the U. S. and has recently become dominant in the rest of the world.

International Organization for Standardization/Open Systems Interconnection (ISO/OSI)

The ISO/OSI protocol suite is overseen by the International Organization for Standardization (ISO) and is known as the reference model for Open Systems Interconnection (OSI). It has been developed by international standards-making bodies since 1980. A component standard of particular importance to information access is the National Information Standards Organization's Z39.50 that is widely referred to as the Information Retrieval Protocol. The Government of the United States has also officially adopted the OSI protocols as the future standard of the government. They have adopted the Government Open Systems Interconnection Protocols (GOSIP). However, since the full OSI protocol suite is still being implemented, it has not been as widely adopted as TCP/IP. It has the strong advantage of being built on a global standards effort and widely accepted as a future protocol suite. As a practical matter, work is now going on to merge the two protocol suites in an approach which will allow transparent interaction among networks based on the two different suites.

---

[3]A fuller discussion of protocols than possible for this paper can be found in John S. Quarterman, *The Matrix, Computer Networks and Conferencing Systems Worldwide, Digital* Press, Bedford, MA, 1990. Quarterman provides a very extensive survey of the computer networking world, and his work is an extraordinarily useful reference.

Layers

Both of TCP/IP and OSI protocol suites have been constructed as layered protocols. Briefly, a layered protocol is made up of standards at each of several levels required for full functionality of the network. For instance, the OSI model divides communications into seven different layers. At the lowest level there is agreement about the physical means of conveying electromagnetic signals from one location to another. This function is provided by the physical layer. At the next level up, the data link and/or network layers implement agreement on the packaging of those signals so the physical connection can be fairly allocated among multiple users, packets can be routed from the source to destination, and at least some degree of error detection provided. Then, at higher levels in the protocol[4], provisions are made to support applications that run on the network. In different protocol suites, these layers may be subdivided differently, but all protocol suites now in use for data networking contain this fundamental concept of layering.

Physical layer

At the physical layer, there are four primary media over which data is transmitted; copper wire, coaxial cable, fiber optics, and radio.

Copper wire. The first, and most heavily used, is plain copper wire. Copper wire forms the foundation for most of our data communications and telephony circuits, especially in carrying information to and from the home and the office to the nearest concentration point or switching site. In the early days of data communications, copper wire was used as an analog medium and was able to transmit only about ten characters per second. The use of microelectronics has enabled copper wire to be used for direct digital transmission and suitably installed copper wire is now capable of millions of characters per second over short distances. Over longer distances, such as those found between a telephone company switching center and a residence or small business, the ISDN Basic Rate service provides speeds in the neighborhood of 18,000 characters per second. Basic Rate ISDN is designed to provide a complete set of communications services including both voice and data. Considerable debate is underway regarding the wisdom of broad deployment of this relatively slow digital transmission technology. As these capabilities demonstrate, copper wire is now capable of providing digital data communication service and has the advantage of already being in place ubiquitously in our institutions, businesses and communities.

Coaxial cable. When higher speed transmission or longer distances must be accommodated, networks can use coaxial cable, a special type of copper wire in which, among other things, the conductors are arranged to resist interference from

---

[4]The higher protocol layers are the transport, session, presentation, and application layers. The transport layer provides for the integrity of the message being transported as opposed to the packets making up that message. The session layer establishes the environment in which a group of messages are combined into a session. The presentation layer is concerned with defining the meaning of the data transmitted, and includes the various codes by which characters are mapped to numbers and more complicated schemes such as that defining the format of an electronic mail message. The application layer provides services to the user, be it a human being or computer. Examples include electronic mail, information bases, etc.

external electrical signals. Coaxial cable can be used in a 'baseband' network to transmit a single stream of digits or it can be used in a 'broadband' network to convey multiple streams of digits along with analog signals such as television and audio. Baseband cable is less expensive to install and maintain but is usually used for only one service at a time. On the other hand, broadband cable is more expensive to install and maintain, but can provide the basis for a large number of services. Coaxial cable, because it is physically larger and requires more complicated connectors, is more expensive to install than regular copper wire and is typically used mainly in cases where its special attributes are required. Advances in optoelectronics have led to a steady decline in coaxial cable installations, with a corresponding increase in the use of fiber optics.

Fiber Optics. The newest means of transmitting information is via pulses of light conducted through glass fibers. This technique, known as fiber optics, has the highest potential for data transmission rates and the lowest susceptibility to error in transmission. Consequently, as the technology for using fiber optics to transmit digital information improves, it is being used for an increasingly broad range of applications in data networking. Fiber optics has been used for some time now to interconnect large telephone switches. More recently, it has been installed in undersea cables for intercontinental transmission of digitized information and is starting to be widely used within institutions and businesses to provide high speed internal data communications. It is now possible to transmit more than 200 million characters per second with production fiber optic technology over substantial distances. Standards have now evolved that allow 10 million characters per second to be routinely transmitted within an institution.

Radio. The final medium for communication is radio. Prior to the rapid increase in cost effectiveness of fiber optics, radio communication was very attractive for long distance service. The two dominant uses today are satellite communications for broad area and international coverage and microwave radio relay towers for connectivity to locations that cannot be economically reached by fiber. As fiber technology has improved, satellite radio is being increasingly relegated to broadcast of information rather than point to point communication or connectivity to remote locations in which it is difficult to justify the installation of optical fiber. For instance, microwave relay is still the dominant means of moving information through the western mountainous segments of the United States, and many transoceanic circuits are still implemented via satellite. There are remote libraries and schools in the north central U. S. and Alaska that use satellite technology because the service density is so low that even microwave relay is prohibitive in cost.

There are now some new initiatives in the use of radio for local area and metropolitan networking that capitalize on radio's wireless nature[5]. Several

---

[5] These services often use "spread-spectrum" technology. It allows multiple stations to simultaneously transmit on a frequency range without the usual problems of allocating separate channels to each transmitter. These systems therefore do not generally need to be licensed as many transmitters can coexist on the same frequency band. Spread spectrum works by assigning a "spreading algorithm" to each transmitter. This algorithm is analogous to a person's individual voice. The receiver then uses another algorithm to decode only the desired sender's code in much the way that people at a party can listen to a specific conversation in a crowded room and perceive every

companies have now released radio-based packet switched local area networks designed for offices where wired networks are either difficult to install (old buildings) or where there is so much movement of the computers that the constant relocation of wiring is a major barrier to network implementation and operation. In addition there are an increasing number of products that provide for connection of personal computers through the switched circuit cellular telephone network, and IBM and Motorola have recently begun providing a nationwide packet radio networking service in support of mobile field employees.[6]

Data Link and Network Layers

A number of different standards are available for the data link and network layers, most of which are specialized to provide support for particular environments. While there are many such protocols, the most notable are:

Ethernet (usually IEEE 802.3) Ethernet is a LAN standard that operates over baseband coaxial cable, broadband coaxial cable, and plain copper wire. Ethernet can support communication at about one million characters per second among several hundred cooperating computers within a relatively small (less than one kilometer end to end) geographic region. Ethernet has for many institutions become a standard for reasonably high speed data networks.

Token Ring (IEEE 802.5) An alternate standard known as the IBM token ring standard was designed primarily to use twisted pair wire and fiber at data rates comparable to Ethernet over similar distances.

Fiber Distributed Data Interface (FDDI) FDDI is a new standard designed for use with optical fiber and can interconnect up to five hundred computers over distances of approximately one hundred miles at speeds of 10 million characters per second. These standards are finalized and equipment able to interface computers to an FDDI network is available. FDDI is being used mainly to interconnect other lower speed networks such as Ethernets and IBM token rings. These connections are made through routers or bridges, small computers that switch packets between LANs, MANs, or WANs.

Integrated Services Digital Network (ISDN) ISDN, mentioned above, is another standard conceived as a significant advance to the existing voice network which is just coming into use. ISDN is designed to operate at a number of speeds. Different protocols and media are used at each of those speeds. The best known implementation of ISDN is the Basic Rate service that is designed to allow for a circuit switched connection capable of carrying digitized voice and data (at about 9000 characters per second) between an office or residence and a telephone company switching center. Several phone companies have now published tariffs for ISDN.

---

other conversation as noise. Wireless LANs can be configured like Ethernet without the expense of wiring but with the flexibility to support mobile workstations. Spread spectrum is also being tested for public wide area data and voice services so that ubiquitous untethered communications can be offered as a service.

[6]The service is called Advanced Radio Data Information Service (ARDIS). It covers 8000 cities and towns in the U. S. and operates at about 500 characters per second. Another service that has been announced and is just becoming operational is Ram Mobile Data.

Several regions of the United States have or will soon have ISDN service available for use as in businesses. It may also be available for residential use in some areas but applications supporting the use of ISDN for residential use are not plentiful.

Primary rate ISDN is also beginning to appear. It allows twenty-three Basic Rate voice or data channels to be combined in one higher speed channel for transmission between switching centers. Much higher rate ISDN service, known as broadband ISDN, is now under development and it will accommodate much higher rates of speed through the use of fiber optics technology.

Synchronous Optical Network (SONET) SONET is a standard for transport of digital traffic over fiber. It is capable of a broad variety of speeds, ranging from approximately 5 million characters per second up to 1 billion characters per second. The SONET technology will provide the foundation for transmission of data in the gigabit (billion bit per second) networks now under discussion.

Asynchronous Transfer Mode (ATM) ATM is a fast packet switching protocol defined to allow various types of traffic to be intermixed over high speed, low error circuits. As such, it is well suited to SONET. Users can gain access to the capabilities of an ATM-based network through protocols such as the IEEE 802.6 metropolitan area service standard or the broad band ISDN CCITT standard.

Switched Metropolitan Data Service (SMDS) SMDS is a service developed by Bellcore to provide public WAN and MAN connectivity. While it is capable of operating across a broad range of speeds, its current trials are being conducted at speeds in the range of 100,000 characters per second.

These physical and network layer technologies can all provide a foundation for higher layers in the protocol stack. In other words, SONET, ATM, and SMDS can provide transport for either TCP/IP or OSI networks. They therefore provide a path for expansion of those protocol families into the wide area network environments provided by optical fiber technology.

Proprietary standards

IBM and Digital Equipment Corporation have also created standards used to support communication among their own products. These standards, which started out as proprietary standards, are now being migrated towards the ISO standard.

System Network Architecture (SNA) IBM's standard, SNA, has been more important to the business community than to higher education and research, but a forerunner of the SNA standard (Network Job Entry) forms the foundation for the BITNET communications network.

DECNET Digital Equipment Corporation's standard, known as DECNET, is moving steadily toward ISO and has provided the foundation for several national and international disciplinary networks.[7]

---

[7]Perhaps the most notable of the international networks is the High Energy Physics Network (HEPnet). Another example is the Space Physics Analysis Network (SPAN), a U. S.-based network with extensions to Japan, Canada, and Europe.

Future Directions

Fortunately, the definition and implementation of the upper layers of the ISO standards is a steadily developing activity. The protocol community has been geographically divided, with the TCP/IP community based primarily in the United States and the ISO community based primarily in Europe. The TCP/IP standard suite has become increasingly popular in Europe, primarily because the European information technology community needs the connectivity offered. As a result, the Internet now provides good access to the European community. The TCP/IP standard now includes definitions at the transport level and below which allow implementation of the upper layers of the ISO suite. There is a steady evolution of ISO applications on top of TCP/IP transport, and the standards appear to be converging. Key vendors are supporting the interoperation of the two dominant suites through development of products that use open data communications standards. If these trends continue, full interconnection of networks and easy passage of information from a network using one set of standards to a network using another will become steadily less cumbersome and more economical.

Applications

All of these systems are justified, of course, by the applications that they enable.

Remote logon

Packet networks were initially expected to be used primarily to support remote use of unique computing resources. For instance, an individual located at one university would be able to remotely logon through the network to a computer at another university that supported applications not locally available. Remote logon continues to be a substantial use of the national network infrastructure and is currently the most important service to the library community because it supports client access to bibliographic data bases. Standards such a Z39.50 and the MARC format for the distribution of catalog records make consistent access to such bibliographic resources possible. There is not yet a widely adopted standard for the user interface to catalogs and other bibliographic information although standards for Common Command Languages have been underway for a number of years.

Electronic mail and conferencing

The second application, not initially anticipated, was interpersonal communication through electronic mail, computer bulletin boards, and electronic conference systems. All of these of systems are in use today and electronic mail in particular is a large contributor to the overall use of the national network infrastructure.

File transfer

From the beginning, it was understood that it would be convenient if data files available on one machine could be easily transmitted to another. This use of the network, or file transfer, is now the largest single use of the national networking infrastructure, accounting for approximately half of the total data transmitted over the network. File transfer will provide the foundation for remote access to published works.

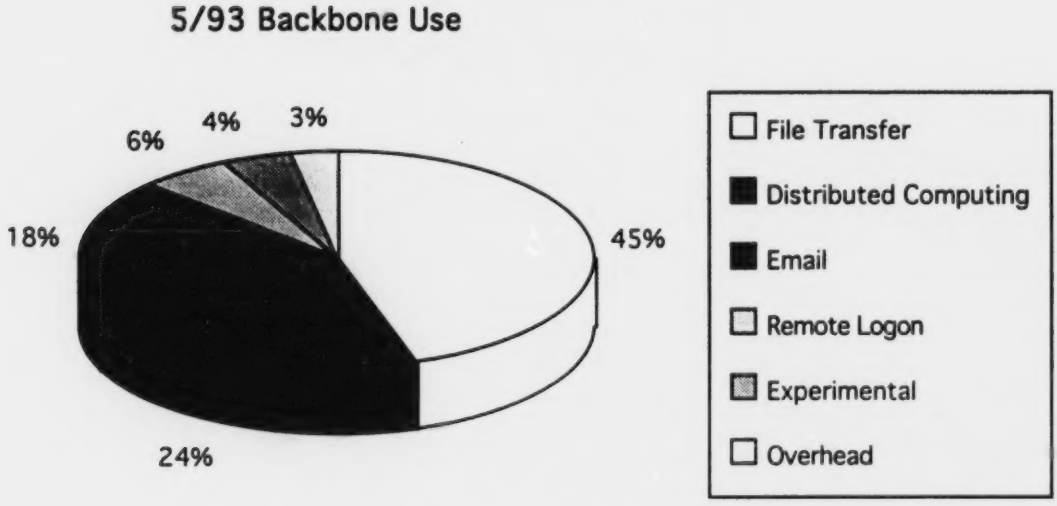

Figure 1

As can be seen in Figure 1, a graph of the percentage of bytes transmitted on the NSFNET backbone during the month of May 1993, remote logon, electronic mail and conferencing, and file transfer comprise almost all of the traffic on the network, with file transfer alone accounting for almost half of the traffic. The remaining fifteen percent is presently not a large share of the backbone traffic, but is expected to grow substantially in coming years.

Client server models

As data networks have increased in speed and the cost of data transmission has steadily declined, another pattern of interaction between computers has become important. These uses are motivated by a desire to combine the virtues of the central computing model that supports shared data access with the economy and ease of use of the personal computing model. The client/server model uses a central computer as a data and/or computation server and a large number of networked personal computers operating as clients. Since the client/server model requires continuous connection between the client machines and the server, and because the responsiveness of the personal computer depends on rapid responses from the server, the model has been implemented mainly in arenas where networking speeds are high and costs are low. The model has found its greatest application in local area networks and is supported with a variety of proprietary software from a number of companies such as Novell, Microsoft, and Banyan.

Increasingly, though, additional efforts are being undertaken to support the client server model in a more standards based environment. In particular, for file sharing, Sun has developed the Network File System which is broadly used in conjunction with the UNIX operating system. Work done at Carnegie Mellon University on the Andrew Project has yielded the Andrew File System which uses server machines running UNIX software communicating over standard TCP/IP networks to client machines of various types in support of file sharing. The Andrew File System is now being extended to even larger environments such as larger university campuses and the national network.

Distributed computing environments

As data networking technology continues to gain in speed and become more cost effective, substantial research and development is underway in support of a fundamentally new type of computing architecture that has been called the distributed computing environment (DCE). In the distributed computing environment, the individual computers of whatever size and capacity are not treated as independent machines but rather as part of an overall system tied together by a comprehensive data architecture and high speed network environment. While such environments have long been the subject of experiment and research, the largest product family to initially achieve commercial success was built by Apollo Computer. As was the case with most computing products of its time, the Apollo distributed computing environment was a proprietary system and lost market share as the open system environments built on standards have become steadily more popular. A number of other manufacturers have since developed similar systems. One category of these systems is built with proprietary software and is directed at personal computers of the MS/DOS or Macintosh class. IBM, Microsoft, and Apple have all brought forward products compatible with their personal computer operating systems to implement proprietary distributed computing environments.

Simultaneously, companies such as Novell and Banyan have been very successful with additional proprietary products that transform the local area networks they sell into distributed computing environments. These environments were initially built to support personal computers in small firms or individual offices and departments in large organizations, but are now being expanded to provide the foundation for distributed computing in very large organizations with geographically dispersed activities.

The other category of distributed computing products is based on open, standards-based software and targeted for the next generation of personal computer and workstation systems, such as UNIX, Windows NT, and OS/2. The Apollo work has been combined with the Carnegie Mellon work on Andrew and been adopted by the Open Software Foundation (OSF) as their standard for distributed computing. Most of the major computer equipment manufacturers are members of OSF, and are now beginning to deliver systems which can be included in a standards-based distributed computing environment. Since these systems support integration of machines and operating systems from many manufacturers, they are particularly well suited for large organizations with heterogeneous computing systems. Further, they support integration of information spaces which span multiple organizations. As a result,

we can expect network based distributed computing environments to become part of the standard products available in the open systems environment over the coming few years.

As the capability and quality of networks continue to grow, DCE will become capable of spanning steadily larger and more geographically dispersed user populations. In the distributed computing environment, remote logon and file transfer are supplanted by automatic access to the computational resources and the files of other computers in the environment. In other words, if the task you are performing at your personal computer requires access to a supercomputer for a particular computation and that computation in turn requires access to a file being constructed on your personal computer, the environment would automatically invoke the supercomputer and connect it to the appropriate file without the need for explicit direction from the user. This type of a computing environment will probably dominate computing toward the end of the decade.

DCE has several characteristics that will determine the network's impact on preservation and access. DCE is based on automatic interaction among a number of geographically distributed file servers, all of which maintain local copies of frequently used files, and only update those files when the master file is changed and local access is desired. Further, if local access is not made, the local copy will be released from storage to be supplanted by other files in more frequent demand. The file server with responsibility for the master copy will have a locally defined system for migrating the document from quickly accessible, expensive storage to less expensive media stored in less expensive space with longer access time if requests for the file decline in frequency. While the file will remain cataloged until explicitly deleted, it may well require hours to physically retrieve and provide network access to a file that hasn't been accessed for a year or more.

Support services

For any of these network based applications to be fully useful, it is important to have available some support services. Support services provide the user with a number of tools and resources that allow him/her to use the networking resources more effectively. The support services are often described by the users of the networking services as the actual service and the networking facilities thus become transparent to the user.

Directory

Perhaps the most obvious support service is the network directory.

"White pages" One might think of this as equivalent to the white pages of the telephone book where you could look up a name and find the number (network address) at which a particular individual or organization could be reached. Because the data networking environment is created by such a large number of organizations working in a loosely linked hierarchy, it is much more difficult to develop and maintain white pages and directory services in the data networking environment than it is in the voice telephone environment. As a result, while there are local examples of user directories, there is no set of comprehensive regional or

national services which provide access to this information. Fortunately, there is an international standards effort which defines standards for such directories and there are experimental implementations of this standard, X.500, now underway at a number of universities and regional networks.

"Yellow pages" Almost equally important to the network user is access to services. There is, for instance, currently no comprehensive directory of bibliographic resources accessible via the network. The same issues that complicate the creation and maintenance of the "white pages" impede the development of a services directory. As the connected community grows, and organizational structures are developed to support fee-for-service delivery over the network, service providers will probably be willing to fund their appearance in the service directory, and it is expected that commercial ventures will build directories and make them available on the network. Once again, there are prototypes and pilots already in existence.

Integrity, security, and privacy

Obviously, as networks proliferate, the issue of information integrity, security, and privacy has become an increasingly important concern. Precisely because of the highly distributed architecture that supports the overall network fabric, much of the information from any particular source to any particular destination traverses computers and packet switches not under the control of either the sending or receiving organization. As a result, it is extremely easy to construct scenarios under which the information is intercepted, inappropriately modified, or prevented from being delivered. While there are very strong norms in the network community regarding the undesirability of such intervention in the network fabric, a substantial amount of the information to be transmitted on the networks is of a nature that the users require more than normative consensus to insure that the network will perform in such a way as to maintain the integrity of the information, secure it from inappropriate tampering, and maintain the privacy of the communications involved. In order to accomplish this, two fundamental technological challenges must be surmounted, encryption and authentication.

Encryption

Encryption is the process by which one ensures that the information itself cannot be viewed by other than the intended recipients. In today's networking environment, the only technique that provides substantial assurance is data encryption before transmission. Ideally, that encryption should be based on well understood international standards for encryption. Unfortunately, the character of the international community has dictated that most of these standards are national as opposed to international. The current standard in the United States is the Data Encryption Standard (DES) and both hardware and software are widely available to implement encryption according to this standard.

User authentication

Successful implementation of encryption depends on a somewhat more difficult service, user authentication. Unless it is possible to definitely ascertain that a user is exactly who she says she is and that a computer being used by that user is in fact

not being inappropriately used by someone else, the best data encryption approaches will not protect the information against users and computers masquerading as authorized to see particular information. Here again substantial work has been undertaken, the most important of which is work done at the Massachusetts Institute of Technology as part of Project Athena to develop the Kerberos user authentication tools. These tools are now being made available in an increasing number of network services, and have been chosen by the Open Software Foundation to be a part of the Distributed Computing Environment. Ultimately, it is likely that user authentication will be integrated with the directory services. With user authentication as a support service, it will be possible to use encryption to secure information being transmitted over the network.

Current Status

Technology

As the international networking environment continues to develop, the technology trends already established seem likely to continue. Protocols are continuing to develop and converge, with a steadily increasing commercial activity growing up to develop and sell the products required to implement and use the standards. The use of the TCP/IP protocols for intracompany networking is accelerating rapidly. Network speeds are increasing. For instance, the NSFNET backbone put its first 45 megabit circuits in service before the end of 1990, and is now providing 45 megabit service throughout the network. A grant to the Corporation for National Research Initiatives from DARPA and NSF is supporting a number of testbed projects through which industry and higher education are building and experimenting with networks capable of transmitting information at speeds greater than 500 megabits. At the same time, the success of the NSFNET effort has generated substantial interest on the part of the major telecommunications vendors, many of which are carrying out technology development efforts in their corporate research and development laboratories.

Experiments are also underway to further develop new uses. In addition to the client-server and distributed computing environment work referred to above, experiments are now underway on the transmission of Group IV facsimile over the internet. Demonstrations have also shown the feasibility of transmitting video and voice on the packet network, although further standards work needs to be done before high quality can be provided over congested networks.

Organization

Distributed management

The evolving international packet data infrastructure is organizationally both highly non-traditional and extremely complex. It may in fact be an exemplar of organizations for our future, but its very novel structure generates interesting issues for both the supporters of the network and the traditional organizations with

which the network must interact. First of all, the network is chiefly held together by a set of shared protocols and values rather than hierarchical management structures and contracts. The protocols and values support a high level of collaborative activity, some of which is governed by legal contracts, but most of which has been carried out in an informal, highly flexible fashion.

It is instructive to trace the path of a packet originating in my home computer in Ann Arbor, Michigan which is bound for a colleague's computer in Ireland. The packet from my computer first enters an Ethernet connecting several different computers in my home. I manage the TCP/IP network connecting those computers as a private network. One of the computers serves as a gateway to an ISDN line provided by Michigan Bell Telephone through a cooperative development project with the University of Michigan. When the packet arrives from the Ethernet to the gateway computer, the address it carries informs the computer that it is next bound for the University of Michigan's campus data network. As a result, the computer signals Michigan Bell Telephone's Ann Arbor central office switch to set up a digital connection from my computer over the ISDN line to a corresponding gateway computer in the University of Michigan's Computing Center. After the Michigan Bell switch responds and establishes the connection, my gateway computer delivers the packet to the University of Michigan gateway computer. That computer, also connected to an Ethernet, sees the address and routes the packet through the Ethernet to the University of Michigan backbone network. From there, the packet travels to the statewide MichNet, which in turn passes it on to the NSFNET backbone network. Once on the NSFNET backbone, the packet travels to a gateway computer in New York which places it on an undersea fiber optic link running to Switzerland where the packets are received by a European supercomputing network. The packet is transmitted over that network via Paris and London to Ireland. Ultimately, the packet reaches a computer on the campus in Dublin, whereupon it is placed on a campus network and delivered to the machine being used by my Irish colleague. Amazingly, all of this activity typically takes place in about a second.

In this process, no fewer than fifteen computers have been involved, owned and operated by at least 10 separate organizations, none of which are under any direct hierarchical management control and most of which can be said to coordinate their activities only in the loosest possible fashion. Yet, the reliability of information delivery through this complicated international web of organizations is in general higher than placing a telephone call or sending a paper letter to the same place. The shared protocols enable this operation, and the shared values of the individuals in all of the participating organizations are responsible for the reliable operation of this complex system.

Not only is this collection of organizations extensive, it is also characterized by agreements ranging from research and development partnerships to contracts for service and not a few handshake agreements. Most of the organizations now providing these linkages did not exist as little as five years ago. One of the significant issues in the planning and further development of this worldwide packet network infrastructure is whether it should continue to develop as a loose confederation of many different types of organizations or whether there should be an effort made to create more of a traditional hierarchical organization controlling

the activities. As different types of organizations with varying objectives provide service in support of the network or attach to the network to allow use, the nature of the required arrangements has become steadily more complex.

For instance, when all of the users of the network were associated with a research organization or an institution of higher education, it was straightforward to make an argument that the infrastructure merited support from federal and state government. Now that the users have a broader range of relationships and the network includes for profit activities such as information provision, it has become less and less clear which part of the infrastructure should be supported by government and which part should be supported by user fees. Within academic organizations, when the use of the facility was primarily the support of research, it was appropriate to think of the financial support coming from overhead on or direct costs of research activities. As the uses expand to include instruction, library access, and facilitation of the operations of the institution, it is less clear that the infrastructure should be supported research funds. In the current financially constrained environment for higher education, various institutions have made differing ad hoc arrangements for network funding.

New providers

In this environment, new organizations are being created to fill in the gaps between the commercial and the not-for-profit community and to provide an avenue for increased funding of the network infrastructure through user fees. Among those recently established are AlterNet, Performance Systems International (PSI), and Advanced Network and Services (ANS), all of which offer TCP/IP packet transport connections and service via facilities they own and connections to the internet. In addition the major telecommunication carriers are beginning to enter the TCP/IP packet transport market, with recent initiatives from Sprint, AT&T, and MCI.

Governmental support

Support from the federal government is also steadily growing. The High Performance Computing and Communications (HPPC) program authorized in 1992 included over $90 million of funding for the development of the national network, all directed at creating a National Research and Education Network (NREN). The HPPC initiative involves the Department of Defense, the Department of Energy, the National Aeronautics and Space Administration, the National Science Foundation, the Department of Commerce, the Environmental Protection Agency, and the National Institutes of Health. The Clinton administration is making a major effort to more rapidly develop the National Information Infrastructure, and the NREN is clearly linked to that effort.

What We Have Learned

Requirements

All of this experience in nationally accessible packet switched networking has

verified three key requirements for the usefulness of the network.

Very wide range of network capacity for each user

First, use of the network by any individual user almost inevitably occurs in bursts of activity. Much of the time the network connection is idle or carrying a very low rate of traffic in support of remote logon to a computer running a character by character application. From time to time, however, such as when a file needs to be moved from one computer to another, or a package of electronic mail needs to be conveyed from one location to another, there is a need to move a comparatively large amount of information in a relatively short time. A researcher may need to conveniently and simultaneously remotely logon to a computer supporting remote visualization of supercomputer results, access a bibliographic database, transmit a revisable document for remote editing, and convey and print the content of an article. As we move into client server models and ultimately to the distributed computing environment, this unevenness of network utilization will increase still further. As systems evolve toward keeping the most-used information close to the user, information will flow over the network in ways not directly related to the user's activity at the workstation. The network supporting this activity must have the instantaneous capacity to run at very high speeds when necessary while incurring minimum costs when demand is very small or non-existent. Packet switching networks are ideal for this sort of use.

Broad addressability reaching user desktop

Second, we have discovered that scholars and researchers need to have nearly simultaneous access to a very broad array of resources. At the same time a researcher is composing electronic mail, (s)he is likely to need to retrieve some information from a database, monitor a large computation going on a remote computer, and have access to a computer bulletin board, all at different network addresses. In order to accomplish this, it is important that the user be able to simultaneously attach to multiple resources through the network. This capability of reaching out for multiple resources enables the user to customize a local computing environment to meet unique needs.

Ties to international industrial and scholarly community

Third, it has become clear that today's researcher or scholar works with an international set of colleagues who themselves are affiliated with organizations not only in higher education but also in industry and government. The network infrastructure needs to provide connections not only within any of these communities, but across these communities. A scholar needs access to commercial databases just as much as (s)he needs access to his or her university's library. A researcher needs access to the supercomputer and data banks at a national laboratory just as much as (s)he needs access to his or her own department's minicomputer. A product engineer needs access to his or her research and development colleagues in the university just as much as (s)he needs access to the parts database upon which his or her product is being built. As a result, the network, while it may require a multiplicity of business and financial arrangements, should provide transparent connectivity across all of the sectors of the economy and providers of information.

Rapid growth underway

Our early experience has also made clear that, even in its comparative infancy, network access is a highly desirable facility. Every standards-based, packet switched network initiative has grown rapidly in the number of computers and users connected and in the amount of traffic it carries. For instance, traffic on the NSFNET backbone has grown more than ten percent per month over the last five years. In fact, the research and scholarly community uses information intensively and extensively and depends heavily on communication with one another to accomplish results.

The need for a broadly cooperative effort

The way in which standards facilitate very broadly distributed investment in support of the creation of a common service has made it possible for funds from an extraordinary variety of sources to be focused on the creation and expansion of the network environment. NSFNET is funded by a combination of educational, industry, and government grants and fees, with the great majority of funding coming from education and industry. Funding from the National Science Foundation is crucial to provide the facility which acts as a stimulus for broad funding, but represents less than ten percent of the total funding of this important resource. On the basis of the experience to date, it is clear that the education and research community is willing to continue to invest substantial resources in expanding both the capacity and breadth of access of the national networking environment.

The future

The higher education and research community is increasingly assuming the NREN's existence built on the experience provided by the NSFNET. There remain, however, a number of difficult policy, funding, and organizational issues.

Policy Issues

As is obvious from the discussion above, successful realization of the National Research and Education Network will depend on resolution of a number of policy issues. While they are all interrelated to some extent, the first four; access, appropriate use, funding, and governance; are quite closely intertwined.

Access

Breadth of access is one of the key determinants of the value of the network infrastructure. As the network reaches a broader community of users, it becomes more valuable to all of its users. On the other hand, the current structure depends to a large extent on a concentration on support of research and education. The network's use, funding, and governance are all evolving in support of a focus on the research and education community.

Appropriate use

If the focus is on the research and education community, what are the appropriate uses? There is consensus that communication among peers within that community is supported, and that services which originate in the community should be available to the community at large. There is developing agreement that services to the community provided from outside the community should also be supported. If those services are not in support of scholarship, however, there is no current agreement on the manner in which they should be provided.

Funding

This is undoubtedly the most troublesome of the policy issues, and is complicated by lack of resolution on the issues of access and appropriate use. If both access and appropriate use were constrained to activities internal to the research and education communities, the current arrangements for shared higher education, industry, and government funding would probably suffice. Since such a limitation would greatly handicap the network's usefulness, a great deal of thought is currently being devoted to providing for fee-for-service types of network use. Such use, it is generally agreed, should pay for its use of network resources, and substantial discussions are underway throughout the networking community to construct the channels through which such funds could flow.

It is clear, even if the network could derive income from fee-based services, that expansion of the network's access beyond the major universities and research laboratories as well as continued rapid advancement of the network's capabilities will continue to require significant government support. Even though that support will continue to constitute a small percentage of the overall support, it is the crucial anchor for investment by the higher education and industrial community.

Governance

The governance issue is particularly complicated. The current, broadly distributed, interlocking set of informal relationships characterizing the network is ideally suited to the community now served by the network and the network's need to change and evolve rapidly. As the network evolves, an increasingly larger constituency needs representation, and feels a need for clarity and structure in governance. The community is actively discussing these issues, but there is no consensus yet visible.

Security

The need for improved network security has ironically generated most of the general public awareness of the network's existence. In fact, the security issues are now mainly confined to the computers attached to the network, and not the network itself. As network support services develop, however, the issues of security will become important to the network itself. Even in the current environment, an increasingly large proportion of the resources required to operate the network are being devoted to control and monitoring of inappropriate use of the network.

Unfortunately, in this arena, as in the privacy and free speech arena, there is not a broad community consensus on appropriate behavior. As the network community develops, these norms will be extremely important to the health of the enterprise.

Privacy and free speech

As always, there are tradeoffs involved in stimulating appropriate use, providing network security, and protecting the privacy and free speech rights of network users. The network already provides access to an extremely broad set of information, some of it offensive to many, and the network is expected to extend its reach. As a result, there has already been extensive coverage of free speech and appropriate use issues in the general press. Recent gatherings of the leaders of the network community have highlighted the concerns over the compromises involved, and increasing concern has been expressed that the governance structure for the network make provision for community involvement in resolving the issues.

Potential Impact

The document life cycle

To assist in understanding the consequences that a ubiquitous national information network could have on preservation and access to scholarly material, it is useful to examine the life cycle of a document and how a networked world might change the way scholars produce and use documents.

Scholars in almost all fields develop and share information in a remarkably similar fashion. The author typically starts a publication by developing an outline of the area she wishes to cover. Using the distributed computing environment supported by the national network, she might well make the outline accessible to a small group of colleagues. The scholar then produces a rough draft of her paper and, at some point, shares that draft with a few close associates and colleagues. Before the advent of the personal computers and the network, the paper was typically a typed script and reproduced via photocopier and facsimile. In the networked world the paper is already in digital form and stored on a file server. The draft will therefore be shared with other scholars by providing them with the network name of document. They will retrieve it, perhaps printing a copy locally, through their own workstation, which will also have access to the file server via the network.

After gaining feedback from colleagues, the scholar will typically present a version of the paper at a scholarly conference, perhaps distributing copies not for attribution. In some cases, the paper will be incorporated into a set of conference proceedings and be published for the first time. While that publication is today accomplished by producing a paper copy of the proceedings for distribution to the attendees, the network will allow the paper to be included in a catalog that references the file server on which the final copy of the paper is maintained. After further collegial comments, the scholar is likely to enhance the paper and submit it to a professional journal in her field for peer review and publication. Usually, by

this time, the paper has been read and the material is familiar to as many as a hundred other scholars working in closely related areas.

Only with its proceedings or journal publication does the paper enter a public bibliography and become available to students and scholars in libraries throughout the world. The delay from the initial creation of the paper to its broad availability is partly due to its iterative refinement. Scholars do not wish to provide their papers for broad distribution until they are confident that they have thought through all of the issues in a careful fashion.

Barriers to scholarship

There are, however, other delays in the process, often as long as a year or more, that are simply a result of time constants built into the current review and journal publication process. The library acquisition and cataloging process can often add additional time delays, especially to information that ultimately finds its way into book form rather than into a scholarly journal.

Unfortunately, libraries are under increasing budgetary pressure that increases the difficulty of providing a comprehensive set of periodical subscriptions and purchasing the full set of scholarly monographs published in all of the fields covered by the modern university. As a result, the process of scholarly publication and distribution of scholarly information are becoming steadily less adequate.

Publishers are also under economic pressures and are finding it very difficult to interrupt this trend. The cost of publishing materials continues to increase and these cost increases must be recovered with steadily smaller publication runs. The specialization of certain publishers in scholarly materials means that there is less and less opportunity for substantial cross-subsidization of scholarly publications by more popular and profitable publications. Furthermore, while the publication process could be organized differently to substantially reduce costs, publishers are prevented by a set of antitrust rulings from collaborating to achieve those new approaches.

Libraries are challenged by the cost of acquiring materials. Recent studies by the Association of Research Libraries found that the price of serials has increased over 50% in a three year period. Many libraries are finding it impossible to keep purchasing all of the serial titles currently in their collection. Rather they are being forced, by the price increases in serials, to make a decision between buying books (monographs) and a reduced number of serials. The Association for Research Libraries found that many of their member libraries were opting to maintain their expenditures on serials at a constant level (thus reducing the number of serials purchased) while reducing the monograph budget. The result is that scholars are handicapped by the less reliable and timely link between the development of new knowledge and its eventual appearance on the library shelf.

The network-enabled model

An electronically facilitated publishing environment could have a positive impact on the price a library would pay for information resources by potentially reducing or eliminating some of the steps currently mandatory in the publication process.

Electronic publishing also has the real potential of increasing the currency of information in the libraries.

Universities, working through their libraries, information technology organizations and, where available, university presses, are likely to be able to provide a new direction for scholarly publication. Most university scholars already produce the early drafts of their publications on personal computers that are increasingly tied together through networks that allow for the transmission of papers within the university and their subsequent sharing among scholars. We are beginning to see the emergence of standards for the production of publication quality information, not only printed text, but also pictorial and graphic information. Modern laser printer technology allows on-demand printing with quite acceptable quality. In other words, the technological structure for electronic publication is already in place in many universities and will shortly be in place in almost all institutions of higher education.

Some modest organizational changes and investments would produce an environment in which individual scholars could use these capabilities to write and publish their scholarly work and acquire the scholarly work of others. Electronic publication mechanisms could be created which would interact directly with the scholar and depend on the library for direct "publication."

For example, when the scholar had refined her paper to the point that she wished to make it available to other scholars that she didn't know, the library could catalog the unpublished paper so it could be found and read by scholars throughout the community. In order not to overburden the cataloging staff, the scholar herself would interact with an expert system that would guide her through the development of a preliminary cataloging record that would be inserted in the library's online catalog with a special notation indicating it had not yet achieved the status of a formal catalog record.

When the scholar decided that she wished to submit the article to a journal, she would send electronic mail to the editor of the journal suggesting that the journal's panel of reviewers review the article. She would, of course, send along an electronic copy of the article. The review panel might well suggest further changes in the article but, assuming that it was finally accepted for publication, the article would simply be included in the table of contents of that electronic journal.

The library would be notified electronically that an addition had been made to the table of contents of the journal. At that point, the informal catalog record developed by the faculty member would be reviewed by the journal and by the library preparatory to inclusion in the formal electronic catalog of the library. The paper would then gain the status of full publication. A similar process would occur in the case of a book publication, but with a publisher (perhaps a university press) replacing the journal review board.

To facilitate this new approach to scholarly information, libraries would install facilities that provide for the printing on demand of electronic publications. Scholars, through their professional societies, would organize review panels that decide whether or not a particular paper already in electronic form meets the

standards of quality required to be admitted to the table of contents of an electronic journal. University presses would distribute electronic versions of books for on-demand printing.

The library would not only avoid many of the costs involved in subscribing to a printed journal or buying a book, but would also become a more integral part of the active scholarly life of the university by capturing publications for broad access at an early stage in their creation.

Higher education, working cooperatively, could create a new model for the publication of scholarly information that would substantially reduce the time lags currently experienced, reduce the costs and thereby increase the coverage and comprehensiveness of accessed information and encourage more cross-disciplinary communication.

Preservation

This document life cycle has a number of implications for preservation and access to library materials. The first, and most obvious, is that digitized information can be preserved essentially forever with adequate attention to periodic copying to new media.[8] Thus, each individual digital copy of a document can be preserved indefinitely. Further, since the network will, in order to promote efficient access, maintain multiple copies of each document in use, the probability of losing all copies is statistically very small in spite of the normal disasters which have destroyed so many valuable items in the past. The steadily declining standards for and cost of digital encoding and transmission of documents is providing an increasingly attractive substitute for physical conveyance of hard copy or microfilm. Finally, since the network will provide every user access to a "perfect" copy, local possession of a copy can be based strictly on frequency of reference rather than on accessibility.

Universities as a human knowledge repository

The most difficult issues will arise regarding documents not in continual use. The system will naturally operate to preserve information for which there is frequent demand. Unfortunately, the system, unless specifically designed for preservation of all materials held, will have a tendency to lose access to unreferenced information. As information is increasingly less used, it will be migrated to storage technologies less and less suited to rapid access. A very small number of master copies could eventually be stored in inactive archives well away from the information processing facilities. Even if the storage is done on archival quality media, the technology for reading and copying the media will have a limited lifetime. As a result, a program that regularly renews digital copies of unreferenced material will be necessary if it is to be preserved in digital form.

While the network will make it feasible for this copying to be concentrated in a few locations with special facilities and funding, this new approach to preservation will require a quite different organization and funding approach. The document will

---

[8]See discussion in *Preservation and Access Technology*,A Report of the Technology Assessment Advisory Committee to the Commission on Preservation and Access, August 1990, pp. 45-46.

have to be thought of in terms of its maintenance cost as well as its acquisition cost. These costs will be incurred only for documents not in use, so the motivation for preservation will not be driven by demand for access in the present, but will have to be thought of as an investment in potential access by future generations. In today's environment, only research libraries have undertaken the responsibility of preservation of little used documents. There is no apparent likelihood of the network and associated digital storage technology relieving higher education of its responsibility in this arena.

Broad knowledge access

Higher education

Smaller institutions of higher education have a special challenge in maintaining library collections sufficient to support the quality of their programs. One very significant advantage of a highly networked environment that includes electronic library resources is that the distinction between the resources held by a large research university and the smaller institution could be minimized.

Smaller institutions often find it difficult to provide laboratory and research environments that meet the demands of high quality faculty. The ability to collaborate electronically has the potential of providing these faculty members with the ability to participate in significant collaborative activities. This also has the potential of increasing the number of scientists, engineers and scholars who actively work on significant projects.

K-12

The educational environment of the United States, and indeed its future workforce, is totally dependent on the quality of the K-12 educational environment. The demands currently placed on the K-12 system are outstripping the ability of this system to meet the additional requirements being placed on it. Meeting those demands by expanding traditional resources, (e.g., more teachers, more classes, more specialization and more materials) is clearly beyond the fiscal means of most school districts. Extension of the network could facilitate increased productivity of the currently available resources and bring higher quality services to the student, teachers, and administrators of these schools. Support from a broad coalition of government agencies, universities, and businesses has allowed substantial pilot use of the network in support of K-12 education. Preliminary results from those pilots show immense potential.

International

The educational and library resources required by scholars, students and others, are not geographically centered in any location. Current restrictions on both people and materials make it particularly difficult to ensure that the best resources are being used. The electronic environment described in this article helps to remove these boundaries and restrictions, particularly if we continue to work diligently for international telecommunications collaboration among educational institutions, library networks, and national libraries.

The university

University mail service, copying, etc.

In many U. S. universities electronic mail has already replaced much of the paper-based internal mail service. The electronic distribution of items that would previously have been copied and sent through the normal mail services have evolved so that the inclusion of spreadsheets and graphics is feasible. Items can, of course, be printed if needed. The majority of items sent electronically are probably never transferred to paper.

The Collaboratory

The ability to communicate and share significant quantities of all types of information is now resulting in significantly changed patterns of behavior. There are instances of papers being jointly created by individuals who are not only geographically separated but who have never met in person. This collaborative behavior is possible because of the existence of networked facilities. This environment in which the collaborative behavior is explicitly enhanced is being referred to as a Collaboratory. Collaboratory experiments have the real potential of increasing the number of high quality scientists working on an experiment or bringing together the best and the brightest to solve a problem in spite of the physical barriers that often preclude such collaboration.

Structure

The underlying structure of the university is based on the interests of its scholars and the needs of the students of the university. The interests of scholars are simultaneously becoming more specialized and interdisciplinary resulting in simultaneous requirements for increased depth and breadth of knowledge. The electronic environment will assist the professor and student alike in assuring that the knowledge base available to that person is vast and with sufficient research, that she has the ability to effectively navigate this vast environment. The resources of the university will thus shift from that of a warehouse of information resources to allowing the available resources to be effectively used. Many of the activities of the university will be fundamentally altered as a result of this paradigm shift. As one example, the ownership of information items will become less important than our ability to effectively and efficiently access and use the information.

Conclusion

The success of the NSFNET and the larger Internet of which it is a leading component has supplied, through the NREN Program, a path to a National Information Infrastructure (NII). The NII will have an immense impact, not only on the library community, but on the fashion in which research, development, scholarship, and teaching are carried out in our society. The extension of these capabilities to the broader society will provide opportunities to transform information provision and the way in which people work together. Libraries and

educational institutions will have the opportunity to play a critical role in the networked future. Extraordinary effort and wisdom will be required to insure that our institutions respond to the opportunity and that information technology is used to realize the promise of a more productive, better informed society.

Appendix -- Brief History

Introduction

There are a variety of networks in operation throughout the nation today. Rather than attempting an exhaustive description[9], those networks most relevant to the library community will be described briefly.

ARPANet

Development of the packet switched network infrastructure has been actively underway since the 1960's. The first substantial implementation was the ARPANet, developed with support from the Department of Defense (DOD) Advanced Research Projects Agency (ARPA). The ARPANet technology led to the development of the TCP/IP protocol suite which is now the defacto standard for the higher education and research community. Access to the ARPANet was initially restricted to government funded researchers, but by the end of its life had in fact been substantially broadened. The ARPANet was managed by the Defense Communication Agency, with operations at Bolt, Baranek, and Newman (BB&N) and a Network Information Center at SRI. It was dismantled in 1989; replaced by NSFNET for non-DOD use and by MILNET for DOD activities.

CSNET

CSNET was established in 1981 with initial grant funding from the National Science Foundation, and has been self-supporting since 1985. It is based on the TCP/IP protocol suite, and has been connected to the ARPANet from its inception. CSNET was developed to support the university computer science community and has been open to research and development efforts in both the higher education and industrial communities. CSNET uses both switched and leased circuits, as well as the X.25 protocol. CSNET was originally operated by the University Corporation for Atmospheric Research (UCAR) with network operation, coordination, and an information center at BB&N. CSNET is now merged with BITNET to form the Corporation for Research and Education Networking (CREN).

BITNET

BITNET was initiated as an informal experiment in information exchange between the City University of New York and Yale University using IBM's Network Job Entry (NJE) protocol. It has since grown to become an international network serving thousands of hosts worldwide. BITNET has extensive gateways to other networks, and is perhaps the most ubiquitously accessible network in the world. BITNET has been self-supporting since its inception and has been used primarily in the higher education community. It is now using TCP/IP for much of its data transport capability, but does not directly support the ARPA protocols. BITNET is operated by CREN with a Network Information Center operated by EDUCOM.

---

[9]For full coverage, see Quarterman, *op. cit.*

NSFNET

NSFNET was established in 1987 by the National Science Foundation to provide connectivity among the six recently established national supercomputing centers and the networks emanating from those centers. It uses the TCP/IP protocols, and was from the beginning transparently connected to ARPANet and CSNET. NSFNET is open to research, education, and development activities in the higher education and industrial communities. NSFNET started using 56 kilobit leased lines, non-commercial Internet Protocol(IP) routers and the existing ARPANet infrastructure. Traffic quickly outgrew the capacity of those facilities, and in 1988 the National Science Foundation awarded a 5 year cooperative agreement for construction and operation of a new backbone to MERIT. MERIT is a not-for-profit state networking corporation of Michigan universities with operations in Ann Arbor, Michigan. MERIT established a partnership with IBM and MCI for computing technology and communications facilities, respectively.

During the same period, NSF also funded a number of regional networks that, along with the connections to the supercomputing centers, were interconnected by the backbone. Since its inception, traffic on the backbone has grown at a compounded rate of more than 10% per month. The NSFNET backbone was originally implemented using T1 (1.5 megabit per second) leased lines, and has now been upgraded to T3 (45 megabit per second) leased lines. NSFNET interconnects more than 1000 universities and colleges, and through links to the broader TCP/IP-based internet provides global access to thousands of networks and hosts. It now also provides ISO-based transport capability.

Library networks

A large number of networks have been developed to support library activities. Many of those networks are better described as organizational rather than technological artifacts. Rather than attempting to characterize the many regional networks serving particular library groupings, those networking initiatives that provide broad, national access to large utilities will be described. The first is operated by the Online Computer Library Center (OCLC) and provides nationwide access to the bibliographic database maintained by OCLC in Columbus, Ohio. The second is operated by the Research Libraries Group (RLG) and provides nationwide access to its owner institutions to a large bibliographic database in Palo Alto, California. While both networks were created on the basis of proprietary protocols, they have agreed to use OSI protocols and have gateways to the TCP/IP internet for limited sets of services. Both of these systems, worked with the Library of Congress and other large bibliographic utilities, on the Linked Systems Project. The goal of this project was to allow the bibliographic assets of these systems to be shared and accessed in a distributed fashion. This project resulted in the creation of the Information Retrieval Protocol, Z39.50. This protocol allows users to search and retrieve bibliographic items from other systems without the user having to understand the command language, bibliographic structure, or system specific characteristics of the remote system. This protocol has now been adopted as a national standard by the National Information Standards Organization (NISO) and the American National Standards Institute (ANSI).

For more information on The Commission on
Preservation and Access and its publications
Call or write -
The Commission on Preservation and Access
1400 16th Street, N.W., Suite 740
Washington, DC 20036-2217
(202) 939-3400

BEST COPY AVAILABLE